

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-268683
 (43)Date of publication of application : 20.09.2002

(51)Int.Cl. G10L 15/28
 G10L 15/00
 G10L 15/22
 G10L 21/02

(21)Application number : 2001-067221
 (22)Date of filing : 09.03.2001

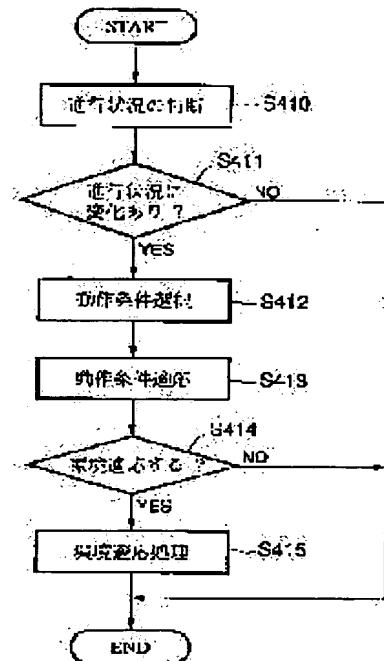
(71)Applicant : CANON INC
 (72)Inventor : YAMAZAKI SHINICHI
 NAKAGAWA KENICHIRO
 YAMAMOTO HIROKI

(54) METHOD AND DEVICE FOR INFORMATION PROCESSING

(57)Abstract:

PROBLEM TO BE SOLVED: To actualize stable voice recognition which is not affected by environmental changes by making it possible to properly set operation conditions of voice recognition processing corresponding to the environmental changes.

SOLUTION: A voice recognition part is provided which inputs a vocal indication to a presentation application being executed. According to words inputted from the voice recognition part, and environmental noise and lightness, the current progress state of an assumed schedule is discriminated (step S410). According to the discriminated progress state, the operation conditions of the voice recognition part are set (steps S412 to S415).



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号
特開2002-268683
(P2002-268683A)

(43)公開日 平成14年9月20日 (2002.9.20)

(51)Int.Cl. ⁷	識別記号	F I	マークコード(参考)
G 10 L 15/28		G 10 L 3/00	5 7 1 D 5 D 0 1 5
15/00			5 5 1 G
15/22			5 7 1 V
21/02		3/02	3 0 1 F

審査請求 未請求 請求項の数38 OL (全 10 頁)

(21)出願番号	特願2001-67221(P2001-67221)	(71)出願人 000001007 キヤノン株式会社 東京都大田区下丸子3丁目30番2号
(22)出願日	平成13年3月9日 (2001.3.9)	(72)発明者 山▲寺▼信一 東京都大田区下丸子3丁目30番2号 キヤ ノン株式会社内
		(72)発明者 中川 寧一郎 東京都大田区下丸子3丁目30番2号 キヤ ノン株式会社内
		(74)代理人 100076428 弁理士 大塚 康徳 (外3名)

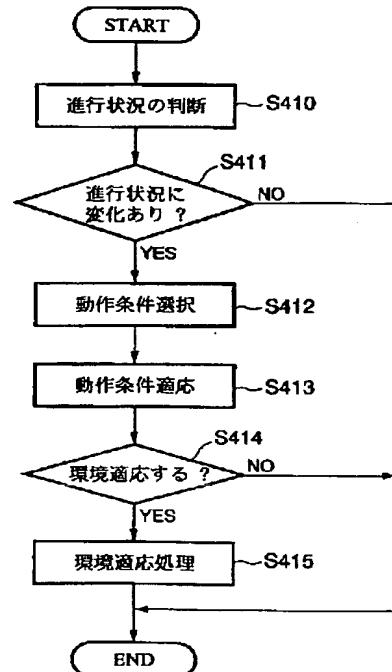
最終頁に続く

(54)【発明の名称】 情報処理方法及び装置

(57)【要約】

【課題】環境の変化に追従して音声認識処理の動作条件を好適に設定することを可能とし、環境の変化に影響されない、安定した音声認識を実現する。

【解決手段】実行中のプレゼンテーションアプリケーションに対して、音声によって指示入を行なうための音声認識部を設ける。この音声認識部から入力される言葉や、環境のノイズ、明るさに基づいて、予め想定されたスケジュールにおける、現在の進行状況が判定される(ステップS410)。そして、判定された進行状況に基づいて音声認識部における動作条件が設定される(ステップS412～S415)。



【特許請求の範囲】

【請求項1】 実行中のアプリケーションに対する指示入力を音声によって行なうための音声認識手段と、予め想定されたスケジュールにおける、現在の進行段階を判定する判定手段と、前記判定手段で判定された進行段階に基づいて前記音声認識手段における動作条件を設定する設定手段とを備えることを特徴とする情報処理装置。

【請求項2】 前記設定手段は、前記音声認識手段で用いる音響モデルの設定、文法あるいは言語モデルの設定、音声を分析する音響分析方法の設定、音響分析を行なうための条件の設定、音声区間検出の閾値の設定、音声認識時の探索条件の設定、音声認識結果の棄却処理に用いる閾値の設定の少なくとも1つを行なうことを特徴とする請求項1に記載の情報処理装置。

【請求項3】 前記スケジュールの進行段階に対応する前記音声認識手段の動作条件を記述した動作条件テーブルを備え、前記設定手段は、前記判定手段で判定した現在の進行段階に対応する動作条件を前記動作条件テーブルより抽出し、該抽出した動作条件に従って前記音声認識手段の動作条件を設定することを特徴とする請求項1または2に記載の情報処理装置。

【請求項4】 前記動作条件テーブルを作成する為のユーザインターフェースを提供する動作条件テーブル作成手段を更に備えることを特徴とする請求項3に記載の情報処理装置。

【請求項5】 前記判定手段は、外部より取り込まれた信号が所定の遷移条件を満足するか否かを判定し、該判定結果に基づいて進行段階を決定することを特徴とする請求項1に記載の情報処理装置。

【請求項6】 前記遷移条件として、前記音声認識手段による認識結果における所定のキーワードの存在を含むことを特徴とする請求項5に記載の情報処理装置。

【請求項7】 環境の光量を検知する検知手段を更に備え、

前記遷移条件として、前記検知手段で検知される光量を含むことを特徴とする請求項5に記載の情報処理装置。

【請求項8】 環境の雑音レベルを計測する計測手段を更に備え、

前記遷移条件として、前記計測手段で計測される雑音レベルを含むことを特徴とする請求項5に記載の情報処理装置。

【請求項9】 前記遷移条件が、操作者によって所定の操作信号が入力されたか否かを含むことを特徴とする請求項5に記載の情報処理装置。

【請求項10】 前記遷移条件として、実行中のアプリケーションの進行状況を含むことを特徴とする請求項5に記載の情報処理装置。

【請求項11】 前記アプリケーションは画像提示機能

を含み、前記アプリケーションの進行状況とは該アプリケーションが提示中の画像であることを特徴とする請求項10に記載の情報処理装置。

【請求項12】 前記スケジュールにおける各進行段階について、前記遷移条件を記載した判定テーブルを更に備え、前記判定手段は、前記判定テーブルを参照して前記進行段階の判定を行なうことを特徴とする請求項5乃至11のいずれかに記載の情報処理装置。

10 【請求項13】 前記判定テーブルにおいて、1つの遷移について複数の遷移条件を記載可能とする特徴とする請求項12に記載の情報処理装置。

【請求項14】 前記判定テーブルを作成するためのユーザーインターフェースを提供するテーブル作成手段を更に備えることを特徴とする請求項12に記載の情報処理装置。

【請求項15】 前記設定手段は、前記判定手段で判定された進行段階に変化があった場合に、前記音声認識手段における動作条件を設定するためのユーザーインター

20 フェースを提供することを特徴とする請求項1に記載の情報処理装置。

【請求項16】 前記判定手段によって判定した進行段階と前記設定手段で設定した動作条件を通知する通知手段を更に備えることを特徴とする請求項1に記載の情報処理装置。

【請求項17】 前記通知手段はその通知内容を表示画面上に表示し、該表示が動作条件の設定画面を兼ねることを特徴とする請求項16に記載の情報処理装置。

30 【請求項18】 前記判定手段で判定された進行段階に基づいて、音声認識の動作条件を環境に適応させる動作環境学習手段を更に備えることを特徴とする請求項1に記載の情報処理装置。

【請求項19】 実行中のアプリケーションに対する指示入力を音声によって行なうための音声認識工程と、予め想定されたスケジュールにおける、現在の進行段階を判定する判定工程と、

前記判定工程で判定された進行段階に基づいて前記音声認識工程における動作条件を設定する設定工程とを備えることを特徴とする情報処理方法。

40 【請求項20】 前記設定工程は、前記音声認識工程で用いる音響モデルの設定、文法あるいは言語モデルの設定、音声を分析する音響分析方法の設定、音響分析を行なうための条件の設定、音声区間検出の閾値の設定、音声認識時の探索条件の設定、音声認識結果の棄却処理に用いる閾値の設定の少なくとも1つを行なうことを特徴とする請求項19に記載の情報処理方法。

【請求項21】 前記スケジュールの進行段階に対応する前記音声認識工程の動作条件を記述した動作条件テーブルを備え、

50 前記設定工程は、前記判定工程で判定した現在の進行段

3

階に対応する動作条件を前記動作条件テーブルより抽出し、該抽出した動作条件に従って前記音声認識工程の動作条件を設定することを特徴とする請求項19または20に記載の情報処理方法。

【請求項22】 前記動作条件テーブルを作成する為のユーザインターフェースを提供する動作条件テーブル作成工程を更に備えることを特徴とする請求項21に記載の情報処理方法。

【請求項23】 前記判定工程は、外部より取り込まれた信号が所定の遷移条件を満足するか否かを判定し、該判定結果に基づいて進行段階を決定することを特徴とする請求項19に記載の情報処理方法。

【請求項24】 前記遷移条件として、前記音声認識工程による認識結果における所定のキーワードの存在を含むことを特徴とする請求項23に記載の情報処理方法。

【請求項25】 環境の光量を検知する検知工程を更に備え、

前記遷移条件として、前記検知工程で検知される光量を含むことを特徴とする請求項23に記載の情報処理方法。

【請求項26】 環境の雑音レベルを計測する計測工程を更に備え、

前記遷移条件として、前記計測工程で計測される雑音レベルを含むことを特徴とする請求項23に記載の情報処理方法。

【請求項27】 前記遷移条件が、操作者によって所定の操作信号が入力されたか否かを含むことを特徴とする請求項23に記載の情報処理方法。

【請求項28】 前記遷移条件として、実行中のアプリケーションの進行状況を含むことを特徴とする請求項23に記載の情報処理方法。

【請求項29】 前記アプリケーションは画像提示機能を含み、前記アプリケーションの進行状況とは該アプリケーションが提示中の画像であることを特徴とする請求項28に記載の情報処理方法。

【請求項30】 前記スケジュールにおける各進行段階について、前記遷移条件を記載した判定テーブルを更に備え、

前記判定工程は、前記判定テーブルを参照して前記進行段階の判定を行なうことを特徴とする請求項23乃至29のいずれかに記載の情報処理方法。

【請求項31】 前記判定テーブルにおいて、1つの遷移について複数の遷移条件を記載可能とすることを特徴とする請求項30に記載の情報処理方法。

【請求項32】 前記判定テーブルを作成するためのユーザインターフェースを提供するテーブル作成工程を更に備えることを特徴とする請求項30に記載の情報処理方法。

【請求項33】 前記設定工程は、前記判定工程で判定された進行段階に変化があった場合に、前記音声認識工

4

程における動作条件を設定するためのユーザーインターフェースを提供することを特徴とする請求項19に記載の情報処理方法。

【請求項34】 前記判定工程によって判定した進行段階と前記設定工程で設定した動作条件を通知する通知工程を更に備えることを特徴とする請求項19に記載の情報処理方法。

【請求項35】 前記通知工程はその通知内容を表示画面上に表示し、該表示が動作条件の設定画面を兼ねることを特徴とする請求項34に記載の情報処理方法。

【請求項36】 前記判定工程で判定された進行段階に基づいて、音声認識の動作条件を環境に適応させる動作環境学習工程を更に備えることを特徴とする請求項1に記載の情報処理方法。

【請求項37】 請求項19乃至36のいずれかに記載の情報処理方法をコンピュータによって実現するためのコンピュータプログラム。

【請求項38】 請求項19乃至36のいずれかに記載の情報処理方法をコンピュータによって実現するためのコンピュータプログラムを格納する記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、音声による操作が可能な情報処理方法及び装置に関し、特にプレゼンテーションのような講演における資料提示等に好適な情報処理方法及び装置に関する。

【0002】

【従来の技術】 従来より、プレゼンテーションにおいては、OHP用シートに印刷したプレゼンテーション資料を発表者が手動で取り替えながらプレゼンテーションを行っていた。また、最近では、パーソナルコンピュータ(PC)機器上のプレゼンテーションソフトでプレゼンテーションデータを作成し、PC機器とプロジェクタを使用してプレゼンテーション資料の提示を行うのが一般的になりつつある。

【0003】 この種のプレゼンテーションソフトはPC機器のキーボード、マウスを用いて操作するのが一般的であるが、PC機器上のプレゼンテーションソフトを音声認識を用いて操作する手法も幾つか提案されている

40 (例えば、特許第02924717号、特開平6-318235号公報)。音声認識を用いた操作の利点は、ページめくり、効果音の出力、アニメーション表示などプレゼンテーションソフトを操作する際に必要となるマウスなどによるポインティング操作、あるいはキーボードの操作を省くことができ、講演者が発表に集中できることである。また、その操作により発表が中断されることもなく、更に必ずしも発表者がPCの前に居る必要がなくなり、発表スタイルに幅ができるという利点もある。

【0004】

50 【発明が解決しようとする課題】 しかしながら、音声認

5

識の精度は発話者の違い、周囲雑音の変化など使用する環境の変化に影響を受けやすく、音響モデル、音響分析方法、音声区間検出の閾値、探索条件などの音声認識の動作条件は、使用する環境によって最適な設定が異なってくる。

【0005】例えば、一般的なプレゼンテーションにおいては、司会者による講演の紹介に始まり、講演者による講演、講演後の質疑応答という流れが一般的である。そして、これらの流れの中で、講演の進行状況によって、音声認識に関わる環境は変化する。例えば、発話を例にとると、講演中の発話者は講演者に限られことが多いが、質疑応答の場面では、講演者以外にも聴講者や司会者が発言する機会が増える。また、音声認識にとって雑音となる会場内のざわめきも講演の進行状況に応じて変化する。例えば、一般的な講演の場合、講演前後は聴講者が互いに話す声などでざわざわとしているが、講演中は聴講者の注意が講演に移り、ざわめきは小さくなる。

【0006】従って、音声認識の精度は、上記のようなプレゼンテーションの流れの中で生じる環境の変化に影響を受けてしまい、安定した操作が行なえなくなる可能性がある。従来より提案されている手法は、音声認識を行ってプレゼンテーションを操作するものではあるが、いずれの提案においてもプレゼンテーション中に音声認識の精度を改善させるために音声認識処理の動作条件を変更するといった提案はない。

【0007】本発明は上記の課題に鑑みてなされたものであり、環境の変化に追従して音声認識処理の動作条件を好適に設定可能とし、環境の変化によらず安定した音声認識を実現する情報処理装置及び方法を提供することを目的とする。

【0008】また、本発明の他の目的は、講演等、予め想定されたスケジュールの進行状況を検出し、この進行状況に応じて想定される環境に適するように、音声認識処理の動作条件を切換え可能とすることにある。

【0009】

【課題を解決するための手段】上記の目的を達成するための本発明による情報処理装置は例えば以下の構成を備える。すなわち、実行中のアプリケーションに対する指示入力を音声によって行なうための音声認識手段と、予め想定されたスケジュールにおける、現在の進行段階を判定する判定手段と、前記判定手段で判定された進行段階に基づいて前記音声認識手段における動作条件を設定する設定手段とを備える。

【0010】また、上記の目的を達成するための本発明による情報処理方法は、実行中のアプリケーションに対する指示入力を音声によって行なうための音声認識工程と、予め想定されたスケジュールにおける、現在の進行段階を判定する判定工程と、前記判定工程で判定された進行段階に基づいて前記音声認識工程における動作条件

6

を設定する設定工程とを備える。

【0011】

【発明の実施の形態】以下、添付の図面を参照して本発明の好適な実施形態を説明する。

【0012】以下の実施形態では、例えばプレゼンテーション等の発表の進行状況を判断し、進行状況に応じて音声認識の動作条件を変更する音声で操作可能なプレゼンテーション装置を説明する。

【0013】

上述のように、プレゼンテーションの進行に伴う環境の変化に対して安定した認識精度を保つためには、環境の変化に応じてその環境に即した動作条件を選択する必要がある。例えば、静かな環境から周囲雑音が大きい環境に移行する場合、一般的な方法として、音響モデルを雑音に適応させたり、スペクトルサブトラクションに代表される雑音除去処理を音響分析時において行ったりする必要がある。従って、プレゼンテーションの進行において、講演会場のざわめき（雑音）に関し、ざわめきが大きくなるような状況では、先に述べたように、音響モデルを雑音適応したり、音響分析方法を雑音環境に適した方法に変更したりすることが望ましい。

【0014】また、例えば、発話者の変化に対しては、講演中は講演者の音声から作成した音響モデル（特定話者モデル）を用い、質疑応答時は講演者以外の音声も認識できるように多数話者の音声から作成した音響モデル（不特定話者モデル）を用いる方法が考えられる。

【0015】以下、本実施形態のプレゼンテーション装置について詳細に説明する。

【0016】図1は、本発明のプレゼンテーション装置の一構成例を示すブロック図である。図1において、プレゼンテーション装置は、ディスプレイやプロジェクタなどの表示装置101、マイクロフォン等で構成され、音声を入力するために用いる音声入力装置102、光量を検出する光量センサ103、マウスなどのポンティングデバイスやキーボードなどを含んで構成される入力装置104、CPUなどの電子データを処理する中央処理装置105、磁気記憶装置やメモリなどの記憶装置106を具備する。

【0017】図2は、本実施形態によるプレゼンテーション装置の機能構成を示すブロック図である。201は制御部であり、本プレゼンテーション装置全体の処理を制御する。202は音声検出部であり、音声区間検出を行う。また、203は音声認識部であり、言語モデル214や音響モデル213を用いて、音声検出部202で検出した音声を認識し、その認識結果を出力する。なお、音響モデル213には、HMMなどの音声認識に用いる音響モデルが格納されており、言語モデル214には、単語辞書など音声認識に用いる言語モデルが格納されている。

【0018】204は進行状況判断部であり、音声認識

部203よりの音声認識結果や光量センサ103からの信号に基づいてプレゼンテーションの進行状況の変化を検出し、進行状況を判断する。なお、進行状況の判断には、プレゼンテーションの進行状況の変化を検出する条件とその条件に対応するプレゼンテーションの進行状況の変化を記述した進行状況判断テーブル211が用いられる。

【0019】205は動作条件選択部であり、プレゼンテーションの進行状況に対応する音声認識の動作条件を選択する。206は動作条件適用部であり、選択した動作条件を音声検出部202や音声認識部203に適用する。なお、動作条件選択部205及び動作条件適用部206は、プレゼンテーションの進行状況に対応する音声認識の動作条件を記述した動作条件テーブル212を参照して動作条件の選択及び適用を実行する。

【0020】207は進行状況判断テーブル作成部であり、進行状況判断テーブルを作成する。208は動作条件テーブル作成部であり、動作条件テーブルを作成する。209は動作環境適応部であり、音響モデルや音声区間検出の閾値などを使用する環境に適応させる。210はプレゼンテーション制御部であり、ページめくりなどのプレゼンテーション操作を制御する。215はプレゼンテーションデータであり、プレゼンテーションにおいて表示装置101に表示したりするデータ等を含む。

【0021】以上の201～210で示される各機能ユニットは、中央処理装置105が記憶装置106に格納された制御プログラムを実行することによって実現される。また、各機能で参照される、211～215で示されるデータは記憶装置106に格納されている。すなわち、以降で説明する一連の処理は、基本的には各機能を記述したプログラムおよび各機能で参照するデータを記憶装置(106)上にロードし、ロードしたプログラムを中央処理装置(105)が実行することによって実現される。

【0022】以上の構成からなる本実施形態の動作をフローチャートに基づき説明する。

【0023】図3は本実施形態による音声認識を用いたプレゼンテーション装置の動作概要を示すフローチャートである。

【0024】まず、ステップS300において、音声認識処理のための動作条件の更新を行なう。これは、進行状況判断部204がプレゼンテーションの進行状況を判断し、動作条件選択部205及び動作条件適用部206が、上記判断された進行状況に応じた動作条件を音声検出部202や音声認識部203に設定するものである。この処理の詳細については、図4A、図4B等を参照して後述する。ステップ301では、プレゼンテーション装置のユーザの発声を、マイクロフォンなどの音声入力装置102を介して取り込み音声検出部201で音声区間の検出を行う。音声検出部201では、音声入力装置

102から入力された音声データを単位時間(例えば0.01秒)毎のブロックに分割し、このブロック毎に音声データが音声区間であるか非音声区間であるか判定している。音声検出部202で音声区間として検出された場合は、ステップS302に進み、取り込んだ音声を音声認識部203で音声認識する。

【0025】ここで、認識結果が、例えば『次のページ』、『最後のページ』などプレゼンテーション装置に指示を与える特定のキーワードである場合には、ステップS303からステップS304へ進み、認識結果にしたがってプレゼンテーション制御部210でプレゼンテーションの制御を行う(ステップ304～ステップ305)。例えば、『次のページ』というキーワードが認識された場合は、次のページのプレゼンテーションデータを表示装置101に表示する制御を行う。また、認識結果が終了を指示している場合は、処理を終了する。一方、ステップ303において、特定のキーワード以外の単語が認識された場合は、ステップS300へ戻り、再び音声認識の動作条件の更新処理、音声区間の検出処理を行なう。

【0026】上述のような動作をする本実施形態のプレゼンテーション装置において、プレゼンテーションの進行状況を検出し、その進行状況に応じて音声認識処理の動作条件を切換える制御を説明する。

【0027】次に、上述したステップS300における、音声認識の動作条件変更処理について説明する。なお、この動作条件変更処理は、進行状況判断部204、動作条件選択部205、動作条件適用部206が、進行状況判断テーブル211及び動作条件テーブル212を参照して行なう。従って、ユーザは進行状況判断テーブル211及び動作条件テーブル212を、プレゼンテーションの実施に先立って設定しておく必要がある。まず、テーブルの設定操作について説明する。

【0028】図4Aは、実施形態によるテーブル設定処理を説明するフローチャートである。入力装置104からテーブル設定処理のための所定の操作が入力され、当該操作が進行状況判断テーブルの作成処理を指示している場合は、ステップS401からステップS402へ進み、進行状況判断テーブル作成部207を起動し、表示装置101に進行状況判断テーブル作成用の画面を表示する。そして、ステップS403において当該作成用画面への各種操作入力を受け付け、進行状況判断テーブル211を作成し、これを記憶装置106に格納する。

【0029】進行状況判断テーブル作成部207で作成される進行状況判断テーブル211には、進行状況の変化を検出する条件、および条件に対応する進行状況が記述される。ここで、進行状況とは、例えば、『講演前』『講演中』『質疑応答』『講演終了』などのプレゼンテーションの経過や状態を表し、ユーザが自由に定義できる。進行状況の変化の検出には、例えば直前の音声認識

の結果、講演会場内の照明の光量の変化、雑音レベルの変化などの情報を用いる。そして、その時点でのプレゼンテーションの進行状況と、上記情報に基づいて、プレゼンテーションの進行状況に邊かがあるかどうかを判断する。

【0030】図5に進行状況判断テーブルの一例を示す。図5は、直前の音声認識部で特定のキーワードを認識した場合に、プレゼンテーションの進行状況の変化を検出する場合の例である。進行状況が『講演前』の状態の時に、講演開始を意味する『開始します』、『始めます』、『スタート』という音声が認識された場合に、プレゼンテーションが開始されると判断して進行状況を『講演中』とすることが記述されている。また、進行状況が『講演中』である場合に、『終了します』、『以上です』という音声が認識された場合に、プレゼンテーションが『質疑応答』の段階に入ったと判断して、進行状況を『質疑応答』とする。なお、図5に示したように、進行状況判断テーブルには、一つの進行状況の変化に対し、進行状況の変化を検出する条件（ここでは認識結果）を複数対応づけられる。

【0031】同様に、図6は、講演会場の照明の明るさの変化によりプレゼンテーションの進行状況を判断する場合の進行状況判断テーブルの例である。講演会場内の照明の明るさ、すなわち光量は、光量センサ103が検出する。この例では、『講演中』の講演会場内の照明を暗くし、逆に『質疑応答』『講演終了』時には、再度照明を明るくする場合を想定している。従って、図6の進行状況判断テーブルでは、現在の進行状況が『講演前』であって、光量が一定値を下回った場合に、進行状況の判断結果を『講演中』とし、現在の進行状況が『講演中』であって光量が一定値を上回った場合に、進行状況の判断結果を『質疑応答』とするよう記述している。

【0032】同様に図7は、雑音レベルの変化により進行状況を判断する場合の進行状況判断テーブルの例である。ここで言う『雑音レベル』は、非音声区間（ユーザが音声を発していない区間）、すなわち、マイクを通して入力される会場のざわめきなど周囲雑音の音の大きさを意味する。この大きさを測るには、（1）音声区間と非音声区間の切り分け、（2）非音声区間の音の大きさの算出の処理が必要になる。（1）の音声区間と非音声区間の切り分けは、音声検出部202の結果に基づいて音声／非音声を切り分ける。（2）の非音声区間の音の大きさは、一般に、振幅やパワーを用いる。この例では、講演中は講演会場内が比較的静かなものに対し、質疑応答時にはややざわつくことを想定している。従って、図7の進行状況判断テーブルでは、現在の進行状況が『講演前』であって雑音レベル所定の閾値Thを下まわった場合には進行状況を『講演中』に切換え、現在の進行状況が『講演中』であって雑音レベルが閾値Th以上になつた場合には進行状況を『質疑応答』に切換えることが示

されている。なお、雑音レベルの検出は音声入力装置102を介して入力された信号に基づいて行なわれる。

【0033】図4Aに戻り、入力装置104からテーブル設定処理のための所定の操作が入力され、当該操作が動作条件テーブルの作成処理を指示している場合は、ステップS401からステップS404を経てステップS405へ進む。ステップS405では、動作条件テーブル作成部208を起動し、表示装置101に動作条件テーブル作成用の画面を表示する。そして、ステップS406において当該作成用画面への各種操作入力を受け付け、動作条件テーブル212を作成し、これを記憶装置106に格納する。

【0034】動作条件テーブル212には、進行状況に対応する音声認識の動作条件を記述する。動作条件の項目としては、例えば、音声認識に使用する音響モデル、音声認識で用いる文法あるいは言語モデル、音響分析方法、音響分析を行う際の各種条件と、音声区間検出に用いる閾値、音声認識を行う際に用いる各種探索条件と、音声認識結果の棄却処理に用いる各種閾値などがある。

【0035】図8に動作条件テーブルの例を示す。図8の動作条件テーブルでは、各進行状況において選択すべき音響モデルと音声区間検出の閾値、並びに環境適応処理の有無を記述している。図8に示したように、動作条件テーブルには一つの進行状況に対し、複数の動作条件について記述できる。

【0036】入力装置104よりテーブル作成処理の終了が指示されるとステップS407より本処理を終了する。

【0037】次にプレゼンテーション中における音声認識処理の動作条件の変更処理、すなわちステップS300の処理について説明する。音声区間の検出を行う前に、上記図4Aの処理によって作成した進行状況判断テーブル211及び動作条件テーブル212を用いて、音声認識の動作条件の設定が実行される。

【0038】まずステップS410において、進行状況判断部204が進行状況判断テーブル211を参照して進行状況を判断する。本実施形態では、図5乃至図7に示した進行状況判断テーブルを用い、現在の進行状況と、音声検出部202及び音声認識部203からの認識結果、光量センサ103からの光量信号及び音声入力装置102からの信号に基づいて得られる雑音レベルに基づいて進行状況を判断する。この判断の結果、進行状況に変化があった場合は、ステップS411からステップS412に進む。ステップS412では、動作条件選択部205において、動作条件テーブル212を参照して、進行状況判断部204の判断結果（進行状況）に対応する音声認識の動作条件を選択する。

【0039】例えば、図5に示した進行状況判断テーブルを用い、進行状況が『講演中』の場合に『以上です』が認識された場合、進行状況判断部204では、進行状況

11

況判断テーブル（図5）に従って進行状況が『講演中』から『質疑応答』に変化したと判断される（ステップ410、S411）。この場合、進行状況が変化しているので、統いて動作条件選択部205で動作条件テーブルを参照して、ステップS410で判断した進行状況『質疑応答』に該当する動作条件を選択する。図8の動作条件テーブルを用いた場合は、不特定話者用の音響モデルを選択し、音声区間検出の閾値がT_h3に変更される。また、同時に環境適応処理を行うという条件を選択する。

【0040】次に、ステップS413へ進み、ステップS412で選択された動作条件に従い、以降で行う音声認識の各種動作条件の変更を行う。この際に、環境適応処理を行う動作条件が選択された場合は、動作環境適応部209において音響モデルの環境適応処理を行い、適応後の音響モデルを記憶装置106に格納する（ステップS414、S415）。ここでいう環境適応処理は、主に音響モデルの変更を伴うものをいい、公知の技術として、CMS、MLLR、PMCなどの手法を指す。そして、ステップS301以降の音声認識処理では新しく適用された動作条件により、音声認識が実行される。

【0041】以上のように、上記実施形態によれば、音声の内容、会場の明るさ、雑音レベルによりプレゼンテーションの進行状況を検出し、その検出された進行状況に従って各進行状況に対して予め割り当てられた音声認識処理の動作条件が設定される。従って、進行状況に応じて適切な音声認識処理を実現でき、音声による安定したプレゼンテーション操作が可能となる。

【0042】なお、上記実施形態では、進行状況判断部204による進行状況の判断に際して図5～図7に示した全てのテーブルを用いるとしたが、その一部を用いるようにしてもよい。例えば、図6に示した進行状況判断テーブルが用いられない場合、光量センサは不要となる。

【0043】また、上記実施形態のステップS410における進行状況判断において、マウスなどのポインティングデバイスの特定の操作やキーボードにおける特定のキー操作が行われた場合に、進行状況の変化を検出して進行状況を判断してもよい。この場合の進行状況判断テーブルの例を図9に示す。この場合、予め進行状況を切換えるためのキー操作を割り当てて、進行状況判断テーブルに登録することになる。

【0044】更に、上記実施形態のステップS410における進行状況判断において、アプリケーション（本例ではプレゼンテーションプログラム）の実行状況に基づいて判断を行うようにしてもよい。例えば、講演に使用するプレゼンテーションデータのうち、特定のデータを表示した時に進行状況の変化を検出することが挙げられる。より具体的には、例えば、プレゼンテーションデータが有する各表示ページに、タイトル部、1ページ…最

12

終ページ等の属性を持たせておき、図10に示す如き進行状況判断テーブルを参照して、表示中のページの属性に基づいて進行状況を判断する。なお、上述した図5～図7及び図9、図10に示すような進行状況判断テーブルのうちの一つのみが用いられてもよいし、複数を組み合わせて用いてもよい。

【0045】また、実施形態のステップS412における動作条件選択においては、動作条件テーブル212を参照して自動で動作条件を選択したが、図11に示すようなウィンドウを表示し、マウスなどのポインティングデバイスあるいはキーボードを用いて講演者が直接、動作条件を選択できるようにしてもよい。この場合、設定画面の表示は進行状況に変化があったことをトリガとしてもよいし所定の操作によって表示を行うようにしてもよい。

【0046】また、上記実施形態のステップS410における進行状況判断において判断された進行状況を講演者に通知してもよい。また、同様に、ステップS412の動作条件選択において選択された動作条件を講演者に通知してもよい。ここで、講演者への通知の方法としては、表示装置101上に判断された進行状況或いは選択された動作条件を表示するようにしてもよい。本装置の機器構成において、複数の表示装置を設け、講演者のみが見ることのできる表示装置に上記通知内容を表示するようにしてもよい。図12に表示装置上に表示した場合のイメージ図を示す。

【0047】また、図12のごとく表示装置上に表示されている進行状況の部分をマウスなどポインティングデバイスで操作することにより、進行状況を選択できるようにもよい。同様に、表示装置上に表示されている動作条件の部分をマウスなどポインティングデバイスで操作することにより、動作条件を選択できるようにもよい。このようにすることで、プレゼンテーション中にキーワードを認識しにくくなったときに、講演者は進行状況あるいは動作条件をプレゼンテーション中でも任意に変更することが可能となる。

【0048】また、上記実施形態において、動作条件テーブルに記述できる動作条件の項目は、図8に示したものに限らず、例えば、音声認識で用いる文法あるいは言語モデルや、音響分析方法（MFCC分析、LPC-MEL分析、Spectral Subtraction、Cepstrum Mean Subtraction）や、音響分析を行う際の各種条件（サンプリング周波数、分析窓幅、メル係数、Spectral Subtractionの係数、パラメータの種類）や、言語探索を行う際の各種条件（詳細な探索を行なうか行なわないか、使用する言語モデルの種類、言語尤度の重み、探索時に用いる処理軽減方法やビーム幅、求める候補数）や、音声認識結果の棄却処理に用いる各種閾値や、話者適応処理の有無など、音声認識の動作に係わる任意の動作条件を記述できるようにしてよいことはいうまでもない。

【0049】以上説明したように本実施形態によれば、自動でプレゼンテーションの進行状況を判断し、進行状況に応じて音声認識の動作条件を変更することにより、プレゼンテーションの進行状況によらず認識精度の高い、音声認識を用いたプレゼンテーション装置を実現できるという効果がある。

【0050】また、本発明の目的は、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（またはCPUやMPU）が記憶媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。

【0051】この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。

【0052】プログラムコードを供給するための記憶媒体としては、例えば、フロッピディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、CD-R、磁気テープ、不揮発性のメモリカード、ROMなどを用いることができる。

【0053】また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているOS（オペレーティングシステム）などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0054】さらに、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込まれた後、そのプログラムコードの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0055】

【発明の効果】以上説明したように、本発明によれば、環境の変化に追従して音声認識処理の動作条件を好適に設定することが可能となり、環境の変化に影響されない、安定した音声認識を実現できる。また、本発明によれば、講演等、予め想定されたスケジュールの進捗状態

を検出し、この進捗状態に応じて想定される環境に適するよう、音声認識処理の動作条件を切換えることが可能となる。

【図面の簡単な説明】

【図1】実施形態によるプレゼンテーション装置の基本構成を示すブロック図である。

【図2】実施形態によるプレゼンテーション装置の機能構成を示すブロック図である。

【図3】実施形態によるプレゼンテーション装置における音声認識を用いたプレゼンテーションの処理手順を示すフローチャートである。

【図4A】実施形態によるプレゼンテーション装置におけるテーブル作成処理を説明するフローチャートである。

【図4B】実施形態によるプレゼンテーション装置におけるプレゼンテーションの進行状況に応じて音声認識の動作条件を変更する場合の処理手順を示すフローチャートである。

【図5】特定のキーワードが認識された場合に進行状況を判断する場合の進行状況判断テーブルのデータ構成例を説明する図である。

【図6】講演会場の照明の変化により進行状況を判断する場合の進行状況判断テーブルのデータ構成例を説明する図である。

【図7】雑音レベルの変化により進行状況を判断する場合の進行状況判断テーブルのデータ構成例を説明する図である。

【図8】動作条件テーブルのデータ構成例を説明する図である。

【図9】マウスなどのポインティングデバイスの特定の操作や特定のキーボード操作が行われた場合に、進行状況を判断する場合の進行状況判断テーブルのデータ構成例を説明する図である。

【図10】講演に使用するプレゼンテーションデータのうち、特定のデータを表示した時に進行状況を判断する場合の進行状況判断テーブルのデータ構成例を説明する図である。

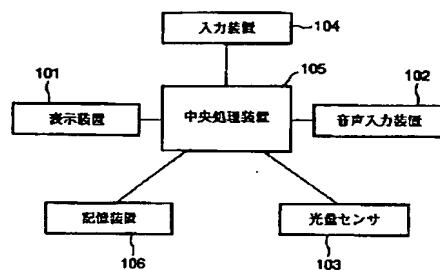
【図11】ポインティングデバイスあるいはキーボードを用いて講演者が直接動作条件を選択する場合の、設定画面例を説明する図である。

【図12】表示装置上に進行状況あるいは動作条件を表示する様子を説明する図である。

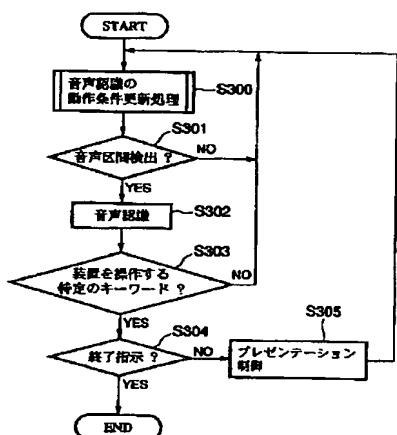
【図8】

動作条件テーブルの例			
進行状況	音響モデル	音声検出の閾値	戻戻適応
講演前	特定話者	Tn1	しない
講演中	特定話者	Tn2	しない
質疑応答	不特定話者	Tn3	する

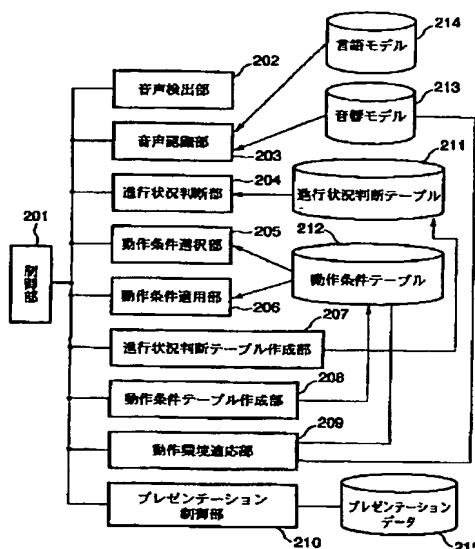
【図1】



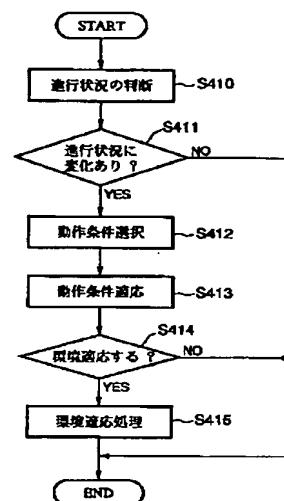
【図3】



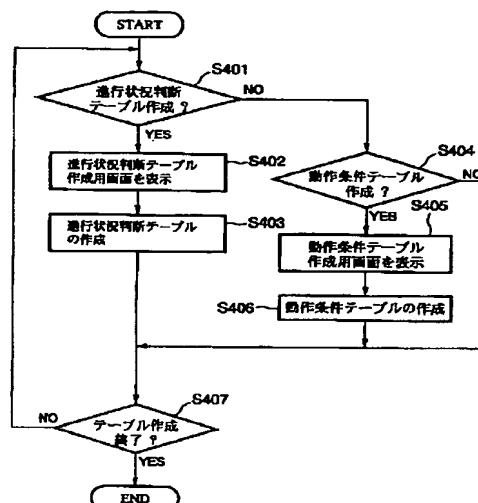
【図2】



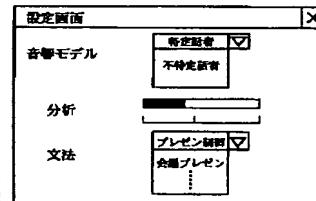
【図4B】



【図4A】



【図11】



【図5】

【図6】

進行状況判断テーブルの例1

条件		進行状況の判断結果
進行状況	認識結果	
講演前	開始します 始めます スタート	講演中 講演中 講演中
講演中	終了します 以上です	質疑応答 質疑応答

進行状況判断テーブルの例2

条件		進行状況の判断結果
進行状況	光量	
講演前	ルクス以下	講演中
講演中	ルクス以上	質疑応答

【図7】

進行状況判断テーブルの例3	
条件	進行状況の判断結果
進行状況 講演前	話者レベル Th 以下
講演中	講演中 質疑応答

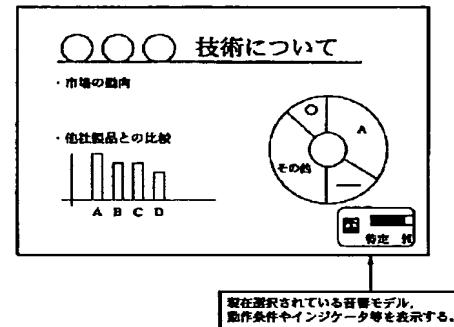
【図9】

進行状況判断テーブルの例4	
条件	進行状況の判断結果
進行状況 講演前	キーボード 改行キーを押す
講演中	講演中 スペースキーを押す 質疑応答
質疑応答	Shift + E キーを押す 講演終了

【図10】

進行状況判断テーブルの例5	
条件	進行状況の判断結果
進行状況 講演前	表示するプレゼンテーションデータ タイトル部表示 1ページ目表示
講演中	まとめ 最終ページ表示 質疑応答

【図12】



フロントページの続き

(72)発明者 山本 寛樹

東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内

F ターム(参考) 5D015 HH23 KK01 LL02 LL03 LL10